



## SEQUENCE LISTING

<110> BioHessan, Inc.  
Woodbury, New York  
deVos, Theo  
Irani, Meher

<120> Recombinant Fusion Proteins with High Affinity Binding to Gold  
and Applications Thereof

<130> Fusionpatent.doc

<140> US 10/671,995  
<141> 2003-09-26

<160> 8

<210> 1  
<211> 5454  
<212> DNA  
<213> Escherichia coli

<220>  
<223> Expression plasmid pPA-GBP, for His6-protein A-GBP. Vector is  
pQE-80L (Qiagen). CDS for Protein A (nucleotides 160-528) is from  
pEZZ18 (Amersham). CDS for GBP (nucleotides 565-858) is from  
pSB3053 from (Brown, Nat. Biotechnol. 15:269-272, 1997).

<220>  
<221> CDS  
<222> (115)..(885)  
<223> CDS for fusion protein His6-protein A-GBP consisting of 257 amino  
acid residues.

<400> 1  
ctcgagaaat cataaaaaat ttatttgctt tgtgagcgga taacaattat aatagattca 60  
  
attgtgagcg gataacaatt tcacacagaa ttcattaaag aggagaaatt aact atg 117  
Met  
1  
  
aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt gcg caa 165  
Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Ala Gln  
5 10 15  
  
cac gat gaa gcc gta gac aac aaa ttc aac aaa gaa caa caa aac gcg 213  
His Asp Glu Ala Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala  
20 25 30  
  
ttc tat gag atc tta cat tta cct aac tta aac gaa gaa caa cga aac 261  
Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn  
35 40 45  
  
gcc ttc atc caa agt tta aaa gat gac cca agc caa agc gct aac ctt 309  
Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu

50	55	60	65	
tta gca gaa gct aaa aag cta aat gat gct cag gcg ccg aaa gta gac				357
Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp				
70		75	80	
aac aaa ttc aac aaa gaa caa caa aac gcg ttc tat gag atc tta cat				405
Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu His				
85		90	95	
tta cct aac tta aac gaa gaa caa cga aac gcc ttc atc caa agt tta				453
Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser Leu				
100		105	110	
aaa gat gac cca agc caa agc gct aac ctt tta gca gaa gct aaa aag				501
Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys Lys				
115		120	125	
cta aat gat gct cag gcg ccg aaa gta gac gcg aat tcg agc tct ggt				549
Leu Asn Asp Ala Gln Ala Pro Lys Val Asp Ala Asn Ser Ser Ser Gly				
130		135	140	145
agt ggc aat ggt cat atg cat gga aaa act cag gca acc agc ggg act				597
Ser Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr				
150		155	160	
atc cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag				645
Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln				
165		170	175	
agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg				693
Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met				
180		185	190	
cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga				741
His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly				
195		200	205	
aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa act				789
Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr				
210		215	220	225
cag gca acc agc ggg act atc cag agc atg cat gga aaa att cag gca				837
Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala				
230		235	240	
acc agc ggg act atc cag agc atg cat gct ctg tcc ctc gag ggt ccg				885
Thr Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly Pro				
245		250	255	
taataagctt aattagctga gcttggactc ctgttgatag atccagtaaat gacctcagaa				945
ctccatctgg atttgttcag aacgctcggt tgccgcggg cgttttttat tggtgagaat				1005
ccaagctagc ttggcgagat tttcaggagc taaggaagct aaaatggaga aaaaaatcac				1065
tggtatatacc accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcattttca				1125

gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg ccttttttaa	1185
gaccgtaaag aaaaataagc acaagtttta tccggccttt attcacattc ttgcccgcct	1245
gatgaatgct catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga	1305
tagtgttcac ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg	1365
gagtgaatac cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg	1425
ttacggtgaa aacctggcct atttcctaa agggtttatt gagaatatgt ttttcgtctc	1485
agccaatccc tgggtgagtt tcaccagttt tgatttaaac gtggccaata tggacaactt	1545
cttcgcccc gttttacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc	1605
gctggcgatt caggttcac atgccgtttg tgatggcttc catgtcggca gaatgcttaa	1665
tgaattacaa cagtactcg atgagtggca gggcgggcg taattttttt aaggcagtta	1725
ttggtgccct taaacgcctg gggtaatgac tctctagctt gaggcacaa ataaaacgaa	1785
aggctcagtc gaaagactgg gcctttcggt ttatctgttg tttgtcggg aacgctctcc	1845
tgagtaggac aaatccgccc tctagattac gtgcagtcga tgataagctg tcaaacatga	1905
gaattgtgcc taatgagtga gctaacttac attaattgcg ttgcgctcac tgcccgttt	1965
ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg	2025
cggtttgct attgggcgcc aggggtggtt ttcttttcac cagtgagacg ggcaacagct	2085
gattgccctt caccgcctgg ccctgagaga gttgcagcaa gcggtccacg ctggtttgcc	2145
ccagcaggcg aaaatcctgt ttgatggtgg ttaacggcg gatataacat gagctgtctt	2205
cggtatcgtc gtatcccact accgagatat ccgcaccaac gcgcagccc gactcggtaa	2265
tggcgcgcat tgcgcccagc gccatctgat cgttggcaac cagcatcgca gtgggaacga	2325
tgccctcatt cagcatttgc atggtttgtt gaaaaccgga catggcactc cagtcgcctt	2385
cccgttcgc tatcggctga atttgattgc gagtgaata tttatgccag ccagccagac	2445
gcagacgcgc cgagacagaa cttaatgggc ccgctaacag cgcgatttgc tggtagacca	2505
atgcgaccag atgctccacg cccagtcgcg taccgtcttc atgggagaaa ataatactgt	2565
tgatgggtgt ctggtcagag acatcaagaa ataacgccg aacattagtg caggcagctt	2625
ccacagcaat ggcaccttg tcatccagcg gatagttaat gatcagcca ctgacgcgtt	2685
gcgcgagaag attgtgcacc gccgctttac aggttcgac gccgcttcgt tctaccatcg	2745
acaccaccac gctggcacc agttgatcgg cgcgagattt aatcgccgcg acaatttgcg	2805

acggcgcggtg	cagggccaga	ctggaggtgg	caacgccaat	cagcaacgac	tgtttgcccc	2865
ccagttgttg	tgccacgcgg	ttgggaatgt	aattcagctc	cgccatcgcc	gcttccactt	2925
tttcccgcgt	tttcgcagaa	acgtggctgg	cctgggtcac	cacgcgggaa	acggtctgat	2985
aagagacacc	ggcatactct	gcgacatcgt	ataacgttac	tggtttcaca	ttcaccaccc	3045
tgaattgact	ctcttccggg	cgctatcatg	ccataccgcg	aaaggttttg	caccattcga	3105
tggtgtcgga	atttcgggca	gcgttgggtc	ctggccacgg	gtgcgcatga	tctagagctg	3165
cctcgcgcg	ttcgggtgatg	acggtgaaaa	cctctgacac	atgcagctcc	cggagacggg	3225
cacagcttgt	ctgtaagcgg	atgccgggag	cagacaagcc	cgtcagggcg	cgtcagcggg	3285
tgttggcggg	tgtcggggcg	cagccatgac	ccagtcacgt	agcgatagcg	gagtgtatac	3345
tggtttaact	atgcggcatc	agagcagatt	gtactgagag	tgccaccatat	gcggtgtgaa	3405
ataccgcaca	gatgcgtaag	gagaaaatac	cgcacagggc	gctcttccgc	ttcctcgctc	3465
actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3525
gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	3585
cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	3645
ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	3705
ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	3765
ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	3825
agctcacgct	gtaggtatct	cagttcgggtg	taggtcggtc	gctccaagct	gggctgtgtg	3885
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	3945
aaccgggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	4005
gcgaggtatg	taggcgggtgc	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	4065
agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	4125
ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtgggttttt	tgtttgcaag	4185
cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	4245
tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	4305
aggatcttca	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	ctaaagtata	4365
tatgagtaaa	cttgggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	4425
atctgtctat	ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	4485
cgggagggct	taccatctgg	ccccagtgc	gcaatgatac	cgcgagaccc	acgctcaccg	4545

```

gctccagatt tatcagcaat aaaccagcca gccggaagg ccgagcgag aagtggtcct 4605
gcaactttat ccgcctccat ccagtctatt aattggtgcc gggaagctag agtaagtagt 4665
tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc 4725
tcgtcgtttg gtatggcttc attcagctcc gggtcccaac gatcaaggcg agttacatga 4785
tcccccatgt tgtgcaaaaa agcgggttagc tccttcggtc ctccgatcgt tgtcagaagt 4845
aagttggccg cagtgttatc actcatgggt atggcagcac tgcataattc tcttactgtc 4905
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 4965
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaa tacgggataa taccgcgcca 5025
catagcagaa ctttaaaagt gtcctcatt ggaaaacgtt cttcggggcg aaaactctca 5085
aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtcacc caactgatct 5145
tcagcatctt ttactttcac cagcgtttct gggtagcaa aaacaggaag gcaaaatgcc 5205
gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cttttttcaa 5265
tattattgaa gcatttatca gggttattgt ctcagagcg gatacatatt tgaatgtatt 5325
tagaaaaata aacaaatagg ggttcgcgc acatttccc gaaaagtgcc acctgacgtc 5385
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 5445
cgtcttcac 5454

```

```

<210> 2
<211> 5448
<212> DNA
<213> Escherichia coli

```

```

<220>
<223> Expression plasmid pStreptavidin-GBP for His6-streptavidin-GBP.
pQE-80L vector(Qiagen). CDS core streptavidin(169-531), (Chilkoti
et al., Proc Natl Acad Sci USA 92:1754-1758, 1995). CDS GBP
(559-852), pSB3053, (Brown, Nat. Biotechnol. 15:269-272, 1997).

```

```

<220>
<221> CDS
<222> (115)..(879)
<223> CDS for fusion protein His6-streptavidin-GBP consisting of 255
amino acid residues.

```

```

<400> 2
ctcgagaaat cataaaaaat ttatttgctt tgtgagcgga taacaattat aatagattca 60

```

attgtgagcg gataacaatt tcacacagaa ttcattaaag aggagaaatt aact atg	117
Met	
1	
aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt ggc cat	165
Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Gly His	
5 10 15	
atg gct gaa gct ggt atc acc ggc acc tgg tac aac cag ctg gga tcc	213
Met Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn Gln Leu Gly Ser	
20 25 30	
acc ttc atc gtt acc gct ggt gct gac ggt gct ctg acc ggt acc tac	261
Thr Phe Ile Val Thr Ala Gly Ala Asp Gly Ala Leu Thr Gly Thr Tyr	
35 40 45	
gaa tcc gct gtt ggt aac gct gaa tct aga tac gtt ctg acc ggt cgt	309
Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val Leu Thr Gly Arg	
50 55 60 65	
tac gac tcc gct ccg gct acc gac ggt tcc gga acc gct ctg ggt tgg	357
Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr Ala Leu Gly Trp	
70 75 80	
acc gtt gct tgg aaa aac aac tac cgt aac gct cac tcc gct acc acc	405
Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His Ser Ala Thr Thr	
85 90 95	
tgg tct ggc cag tac gtt ggt ggt gct gaa gct cgt atc aac acc cag	453
Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu Ala Arg Ile Asn Thr Gln	
100 105 110	
tgg ttg ttg acc tcc ggc acc acc gaa gct aac gcg tgg aaa tcc acc	501
Trp Leu Leu Thr Ser Gly Thr Thr Glu Ala Asn Ala Trp Lys Ser Thr	
115 120 125	
ctg gtt ggt cac gac acc ttc acc aaa gtt tcg agc tct ggt agt ggc	549
Leu Val Gly His Asp Thr Phe Thr Lys Val Ser Ser Ser Gly Ser Gly	
130 135 140 145	
aat ggt cat atg cat gga aaa act cag gca acc agc ggg act atc cag	597
Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln	
150 155 160	
agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg	645
Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met	
165 170 175	
cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga	693
His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly	
180 185 190	
aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa act	741
Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr	
195 200 205	
cag gca acc agc ggg act atc cag agc atg cat gga aaa act cag gca	789

Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala	210	215	220	225	
acc agc ggg act atc cag agc atg cat gga aaa att cag gca acc agc					837
Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala Thr Ser		230	235	240	
ggg act atc cag agc atg cat gct ctg tcc ctc gag ggt ccg					879
Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly Pro		245	250	255	
taataagctt aattagctga gcttggactc ctgttgatag atccagtaat gacctcagaa					939
ctccatctgg atttgttcag aacgctcggg tgccgcgagg cgttttttat tggtgagaat					999
ccaagctagc ttggcgagat tttcaggagc taaggaagct aaaatggaga aaaaaatcac					1059
tggatatacc accgttgata tatcccaatg gcatcgtaaa gaacattttg aggcatttca					1119
gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg cttttttaaa					1179
gaccgtaaag aaaaataagc acaagtttta tccggccttt attcacattc ttgccgcct					1239
gatgaatgct catccggaat ttcgtatggc aatgaaagac ggtgagctgg tgatatggga					1299
tagtgttcac ccttgttaca ccgttttcca tgagcaaact gaaacgtttt catcgctctg					1359
gagtgaatac cacgacgatt tccggcagtt tctacacata tattcgcaag atgtggcgtg					1419
ttacggtgaa aacctggcct atttcacctaa agggttttatt gagaatatgt ttttcgtctc					1479
agccaatccc tgggtgagtt tcaccagttt tgattttaaac gtggccaata tggacaactt					1539
cttcgcccc gttttcacca tgggcaaata ttatacgcaa ggcgacaagg tgctgatgcc					1599
gctggcgatt caggttcac atgccgtttg tgatggcttc catgtcggca gaatgcttaa					1659
tgaattacaa cagtactgcg atgagtggca gggcggggcg taattttttt aaggcagtta					1719
ttggtgccct taaacgcctg gggtaatgac tctctagctt gaggcacaa ataaaacgaa					1779
aggctcagtc gaaagactgg gcctttcggt ttatctgttg tttgtcggtg aacgctctcc					1839
tgagtaggac aaatccgccc tctagattac gtgcagtcga tgataagctg tcaaacatga					1899
gaattgtgcc taatgagtga gctaacttac attaattgcg ttgcgctcac tgcccgttt					1959
ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg					2019
cggtttgctg attgggcgcc agggtggttt ttcttttcac cagtgcgacg ggcaacagct					2079
gattgccctt caccgcctgg ccctgagaga gttgcagcaa gcggtccacg ctggtttgcc					2139
ccagcaggcg aaaatcctgt ttgatggtgg ttaacggcgg gatataacat gagctgtctt					2199
cggtatcgtc gtatcccact accgagatat ccgcaccaac gcgcagcccg gactcggtaa					2259

tggcgcgcat	tgcgcccagc	gccatctgat	cgttggcaac	cagcatcgca	gtgggaacga	2319
tgccttcatt	cagcatttgc	atggtttgtt	gaaaaccgga	catggcactc	cagtcgcctt	2379
cccgttccgc	tatcggctga	atttgattgc	gagtgagata	tttatgccag	ccagccagac	2439
gcagacgcgc	cgagacagaa	cttaatgggc	ccgctaacag	cgcgatttgc	tggtgacca	2499
atgcgaccag	atgctccacg	cccagtcgcg	taccgtcttc	atgggagaaa	ataatactgt	2559
tgatgggtgt	ctggtcagag	acatcaagaa	ataacgccgg	aacattagtg	caggcagctt	2619
ccacagcaat	ggcatcctgg	tcatccagcg	gatagttaat	gatcagccca	ctgacgcgtt	2679
gcgcgagaag	attgtgcacc	gccgctttac	aggcttcgac	gccgcttcgt	tctaccatcg	2739
acaccaccac	gctggcaccc	agttgatcgg	cgcgagattt	aatcgccgcg	acaatttgcg	2799
acggcgcggtg	cagggccaga	ctggaggtgg	caacgccaat	cagcaacgac	tgtttgcccg	2859
ccagttgttg	tgccacgcgg	ttgggaatgt	aattcagctc	cgccatcgcc	gcttccactt	2919
tttcccgcg	tttcgcagaa	acgtggctgg	cctggttcac	cacgcgggaa	acggtctgat	2979
aagagacacc	ggcatactct	gcgacatcgt	ataacgttac	tggtttcaca	ttcaccaccc	3039
tgaattgact	ctcttccggg	cgctatcatg	ccataccgcg	aaaggttttg	caccattcga	3099
tggtgtcggg	atttcgggca	gcgttgggtc	ctggccacgg	gtgcgcatga	tctagagctg	3159
cctcgcgcg	ttcgggtgatg	acggtgaaaa	cctctgacac	atgcagctcc	cggagacggg	3219
cacagcttgt	ctgtaagcgg	atgccgggag	cagacaagcc	cgtcagggcg	cgtcagcggg	3279
tgttggcggg	tgtcggggcg	cagccatgac	ccagtcacgt	agcgatagcg	gagtgatac	3339
tggtttaact	atgcggcatc	agagcagatt	gtactgagag	tgcaccatat	gcggtgtgaa	3399
ataccgcaca	gatgcgtaag	gagaaaatac	cgcatcaggc	gctcttccgc	ttcctcgctc	3459
actgactcgc	tgcgctcggg	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3519
gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	3579
cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	3639
cccctgacg	agcatcaca	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	3699
ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	3759
ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	3819
agctcacgct	gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	3879
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	3939
aaccgggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	3999



gcgaggtatg taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact	4059
agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt	4119
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtgggtttttt tgtttgcaag	4179
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg	4239
tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgag attatcaaaa	4299
aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaataaat ctaaagtata	4359
tatgagtaaa cttggtctga cagttaccaa tgcttaataca gtgaggcacc tatctcagcg	4419
atctgtctat ttcgttcac catagttgcc tgactccccg tcgtgtagat aactacgata	4479
cgggagggct taccatctgg cccagtgct gcaatgatac cgcgagacc acgctcaccg	4539
gctccagatt tatcagcaat aaaccagcca gccggaagg ccgagcgag aagtggctct	4599
gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt	4659
tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt ggtgtcacgc	4719
tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga	4779
tccccatgt tgtgcaaaaa agcggtagc tccttcggtc ctccgatcgt tgtcagaagt	4839
aagttggccg cagtgttatc actcatgggt atggcagcac tgcataattc tcttactgtc	4899
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa	4959
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtaa tacgggataa taccgcgcca	5019
catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcgggcg aaaactctca	5079
aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc caactgatct	5139
tcagcatctt ttactttcac cagcgtttct gggtagcaa aaacaggaag gcaaaatgcc	5199
gcaaaaaagg gaataagggc gacacgaaa tgttgaatac tcatactctt ctttttcaa	5259
tattattgaa gcatttatca gggttattgt ctcagtagcg gatacatatt tgaatgtatt	5319
tagaaaaata aacaaatagg ggttccgcgc acatttccc gaaaagtgcc acctgacgtc	5379
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt	5439
cgtcttcac	5448

<210> 3  
 <211> 1182  
 <212> DNA

<213> Escherichia coli

<220>

<223> Sequence shows CDS portion (115-1293) of the expression plasmid, pPA-GBP-PA for fusion protein His6-protein A-GBP-protein A. The total length of the plasmid is 5859 base pairs. The origin of the basic vector and other CDSs is the same as pPA-GBP.

<220>

<221> CDS

<222> (1)..(1179)

<223> CDS for fusion protein His6-protein A-GBP-protein A consisting of 393 amino acid residues.

<400> 3

atg aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt gcg	48
Met Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Ala	
1 5 10 15	
caa cac gat gaa gcc gta gac aac aaa ttc aac aaa gaa caa caa aac	96
Gln His Asp Glu Ala Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn	
20 25 30	
gcg ttc tat gag atc tta cat tta cct aac tta aac gaa gaa caa cga	144
Ala Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg	
35 40 45	
aac gcc ttc atc caa agt tta aaa gat gac cca agc caa agc gct aac	192
Asn Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn	
50 55 60	
ctt tta gca gaa gct aaa aag cta aat gat gct cag gcg ccg aaa gta	240
Leu Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys Val	
65 70 75 80	
gac aac aaa ttc aac aaa gaa caa caa aac gcg ttc tat gag atc tta	288
Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu	
85 90 95	
cat tta cct aac tta aac gaa gaa caa cga aac gcc ttc atc caa agt	336
His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser	
100 105 110	
tta aaa gat gac cca agc caa agc gct aac ctt tta gca gaa gct aaa	384
Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys	
115 120 125	
aag cta aat gat gct cag gcg ccg aaa gta gac gcg aat tcg agc tct	432
Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp Ala Asn Ser Ser Ser	
130 135 140	
ggt agt ggc aat ggt cat atg cat gga aaa act cag gca acc agc ggg	480
Gly Ser Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly	
145 150 155 160	

act atc cag agc atg cat gga aaa act cag gca acc agc ggg act atc Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile 165 170 175	528
cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser 180 185 190	576
atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His 195 200 205	624
gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys 210 215 220	672
act cag gca acc agc ggg act atc cag agc atg cat gga aaa att cag Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln 225 230 235 240	720
gca acc agc ggg act atc cag agc atg cat gct ctg tcc ctc gag ggt Ala Thr Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly 245 250 255	768
ggc gga tcc ggt tct ggt gcg caa cac gat gaa gcc gta gac aac aaa Gly Gly Ser Gly Ser Gly Ala Gln His Asp Glu Ala Val Asp Asn Lys 260 265 270	816
ttc aac aaa gaa caa caa aac gcg ttc tat gag atc tta cat tta cct Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu His Leu Pro 275 280 285	864
aac tta aac gaa gaa caa cga aac gcc ttc atc caa agt tta aaa gat Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser Leu Lys Asp 290 295 300	912
gac cca agc caa agc gct aac ctt tta gca gaa gct aaa aag cta aat Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys Lys Leu Asn 305 310 315 320	960
gat gct cag gcg ccg aaa gta gac aac aaa ttc aac aaa gaa caa caa Asp Ala Gln Ala Pro Lys Val Asp Asn Lys Phe Asn Lys Glu Gln Gln 325 330 335	1008
aac gcg ttc tat gag atc tta cat tta cct aac tta aac gaa gaa caa Asn Ala Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln 340 345 350	1056
cga aac gcc ttc atc caa agt tta aaa gat gac cca agc caa agc gct Arg Asn Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala 355 360 365	1104
aac ctt tta gca gaa gct aaa aag cta aat gat gct cag gcg ccg aaa Asn Leu Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys 370 375 380	1152
gta gac gcg aat tcg agc tct ggt ggc taa	1182

Val Asp Ala Asn Ser Ser Ser Gly Gly  
385 390

<210> 4  
<211> 1170  
<212> DNA  
<213> Escherichia coli

<220>  
<223> CDS portion (115-1281) of the expression plasmid,  
pStrept-GBP-Strept (total length 5833 base pairs) for fusion  
protein His6-streptavidin-GBP-streptavidin. Vector and other CDSs  
are the same as pPA-GBP and pStreptavidin-GBP.

<220>  
<221> CDS  
<222> (1)..(1167)  
<223> CDS for fusion protein His6-streptavidin-GBP-streptavidin  
consisting of 389 amino acid residues.

<400> 4  
atg aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt ggc 48  
Met Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Gly  
1 5 10 15  
  
cat atg gct gaa gct ggt atc acc ggc acc tgg tac aac cag ctg gga 96  
His Met Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn Gln Leu Gly  
20 25 30  
  
tcc acc ttc atc gtt acc gct ggt gct gac ggt gct ctg acc ggt acc 144  
Ser Thr Phe Ile Val Thr Ala Gly Ala Asp Gly Ala Leu Thr Gly Thr  
35 40 45  
  
tac gaa tcc gct gtt ggt aac gct gaa tct aga tac gtt ctg acc ggt 192  
Tyr Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val Leu Thr Gly  
50 55 60  
  
cgt tac gac tcc gct ccg gct acc gac ggt tcc gga acc gct ctg ggt 240  
Arg Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr Ala Leu Gly  
65 70 75 80  
  
tgg acc gtt gct tgg aaa aac aac tac cgt aac gct cac tcc gct acc 288  
Trp Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His Ser Ala Thr  
85 90 95  
  
acc tgg tct ggc cag tac gtt ggt ggt gct gaa gct cgt atc aac acc 336  
Thr Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu Ala Arg Ile Asn Thr  
100 105 110  
  
cag tgg ttg ttg acc tcc ggc acc acc gaa gct aac gcg tgg aaa tcc 384  
Gln Trp Leu Leu Thr Ser Gly Thr Thr Glu Ala Asn Ala Trp Lys Ser  
115 120 125

acc ctg gtt ggt cac gac acc ttc acc aaa gtt tcg agc tct ggt agt	432
Thr Leu Val Gly His Asp Thr Phe Thr Lys Val Ser Ser Ser Gly Ser	
130 135 140	
ggc aat ggt cat atg cat gga aaa act cag gca acc agc ggg act atc	480
Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile	
145 150 155 160	
cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc	528
Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser	
165 170 175	
atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat	576
Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His	
180 185 190	
gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa	624
Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys	
195 200 205	
act cag gca acc agc ggg act atc cag agc atg cat gga aaa act cag	672
Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln	
210 215 220	
gca acc agc ggg act atc cag agc atg cat gga aaa att cag gca acc	720
Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala Thr	
225 230 235 240	
agc ggg act atc cag agc atg cat gct ctg tcc ctc gag gga tct ggt	768
Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly Ser Gly	
245 250 255	
tct ggt ggc cat atg gct gaa gct ggt atc acc ggc acc tgg tac aac	816
Ser Gly Gly His Met Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn	
260 265 270	
cag ctg gga tcc acc ttc atc gtt acc gct ggt gct gac ggt gct ctg	864
Gln Leu Gly Ser Thr Phe Ile Val Thr Ala Gly Ala Asp Gly Ala Leu	
275 280 285	
acc ggt acc tac gaa tcc gct gtt ggt aac gct gaa tct aga tac gtt	912
Thr Gly Thr Tyr Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val	
290 295 300	
ctg acc ggt cgt tac gac tcc gct ccg gct acc gac ggt tcc gga acc	960
Leu Thr Gly Arg Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr	
305 310 315 320	
gct ctg ggt tgg acc gtt gct tgg aaa aac aac tac cgt aac gct cac	1008
Ala Leu Gly Trp Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His	
325 330 335	
tcc gct acc acc tgg tct ggc cag tac gtt ggt ggt gct gaa gct cgt	1056
Ser Ala Thr Thr Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu Ala Arg	
340 345 350	
atc aac acc cag tgg ttg ttg acc tcc ggc acc acc gaa gct aac gcg	1104

Ile	Asn	Thr	Gln	Trp	Leu	Leu	Thr	Ser	Gly	Thr	Thr	Glu	Ala	Asn	Ala		
	355						360					365					
tgg	aaa	tcc	acc	ctg	ggt	ggt	cac	gac	acc	ttc	acc	aaa	ggt	tcg	agc		1152
Trp	Lys	Ser	Thr	Leu	Val	Gly	His	Asp	Thr	Phe	Thr	Lys	Val	Ser	Ser		
	370					375					380						
tca	agc	tta	att	agc	tga												1170
Ser	Ser	Leu	Ile	Ser													
	385																

<210> 5  
 <211> 1176  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <223> CDS portion (115-1287) of the expression plasmid, pPA-GBP-Streptavidin (total length 5839 base pairs) for fusion protein His6-PA-GBP-streptavidin. Vector and other CDSs are the same as pPA-GBP and pStreptavidin-GBP.

<220>  
 <221> CDS  
 <222> (1)..(1173)  
 <223> CDS for fusion protein His6-PA-GBP-streptavidin consisting of 391 amino acid residues.

<400>	5																
atg	aga	gga	tcg	cat	cac	cat	cac	cat	cac	gga	tcc	ggt	tct	ggt	gcg		48
Met	Arg	Gly	Ser	His	His	His	His	His	His	Gly	Ser	Gly	Ser	Gly	Ala		
1				5					10					15			
caa	cac	gat	gaa	gcc	gta	gac	aac	aaa	ttc	aac	aaa	gaa	caa	caa	aac		96
Gln	His	Asp	Glu	Ala	Val	Asp	Asn	Lys	Phe	Asn	Lys	Glu	Gln	Gln	Asn		
			20					25					30				
gcg	ttc	tat	gag	atc	tta	cat	tta	cct	aac	tta	aac	gaa	gaa	caa	cga		144
Ala	Phe	Tyr	Glu	Ile	Leu	His	Leu	Pro	Asn	Leu	Asn	Glu	Glu	Gln	Arg		
		35					40					45					
aac	gcc	ttc	atc	caa	agt	tta	aaa	gat	gac	cca	agc	caa	agc	gct	aac		192
Asn	Ala	Phe	Ile	Gln	Ser	Leu	Lys	Asp	Asp	Pro	Ser	Gln	Ser	Ala	Asn		
	50					55					60						
ctt	tta	gca	gaa	gct	aaa	aag	cta	aat	gat	gct	cag	gcg	ccg	aaa	gta		240
Leu	Leu	Ala	Glu	Ala	Lys	Lys	Leu	Asn	Asp	Ala	Gln	Ala	Pro	Lys	Val		
65					70				75					80			
gac	aac	aaa	ttc	aac	aaa	gaa	caa	caa	aac	gcg	ttc	tat	gag	atc	tta		288
Asp	Asn	Lys	Phe	Asn	Lys	Glu	Gln	Gln	Asn	Ala	Phe	Tyr	Glu	Ile	Leu		
				85					90					95			

cat tta cct aac tta aac gaa gaa caa cga aac gcc ttc atc caa agt His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser 100 105 110	336
tta aaa gat gac cca agc caa agc gct aac ctt tta gca gaa gct aaa Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys 115 120 125	384
aag cta aat gat gct cag gcg ccg aaa gta gac gcg aat tcg agc tct Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp Ala Asn Ser Ser Ser 130 135 140	432
ggc agt ggc aat ggt cat atg cat gga aaa act cag gca acc agc ggg Gly Ser Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly 145 150 155 160	480
act atc cag agc atg cat gga aaa act cag gca acc agc ggg act atc Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile 165 170 175	528
cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser 180 185 190	576
atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His 195 200 205	624
gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys 210 215 220	672
act cag gca acc agc ggg act atc cag agc atg cat gga aaa att cag Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln 225 230 235 240	720
gca acc agc ggg act atc cag agc atg cat gct ctg tcc ctc gag gga Ala Thr Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly 245 250 255	768
tct ggt tct ggt ggc cat atg gct gaa gct ggt atc acc ggc acc tgg Ser Gly Ser Gly Gly His Met Ala Glu Ala Gly Ile Thr Gly Thr Trp 260 265 270	816
tac aac cag ctg gga tcc acc ttc atc gtt acc gct ggt gct gac ggt Tyr Asn Gln Leu Gly Ser Thr Phe Ile Val Thr Ala Gly Ala Asp Gly 275 280 285	864
gct ctg acc ggt acc tac gaa tcc gct gtt ggt aac gct gaa tct aga Ala Leu Thr Gly Thr Tyr Glu Ser Ala Val Gly Asn Ala Glu Ser Arg 290 295 300	912
tac gtt ctg acc ggt cgt tac gac tcc gct ccg gct acc gac ggt tcc Tyr Val Leu Thr Gly Arg Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser 305 310 315 320	960

gga acc gct ctg ggt tgg acc gtt gct tgg aaa aac aac tac cgt aac	1008
Gly Thr Ala Leu Gly Trp Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn	
325 330 335	
gct cac tcc gct acc acc tgg tct ggc cag tac gtt ggt ggt gct gaa	1056
Ala His Ser Ala Thr Thr Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu	
340 345 350	
gct cgt atc aac acc cag tgg ttg ttg acc tcc ggc acc acc gaa gct	1104
Ala Arg Ile Asn Thr Gln Trp Leu Leu Thr Ser Gly Thr Thr Glu Ala	
355 360 365	
aac gcg tgg aaa tcc acc ctg gtt ggt cac gac acc ttc acc aaa gtt	1152
Asn Ala Trp Lys Ser Thr Leu Val Gly His Asp Thr Phe Thr Lys Val	
370 375 380	
tcg agc tca agc tta att agc tga	1176
Ser Ser Ser Ser Leu Ile Ser	
385 390	

<210> 6  
 <211> 1176  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <223> CDS portion (115-1287) of the expression plasmid,  
 pStreptavidin-GBP-PA (total length 5853 base pairs) for fusion  
 protein His6-streptavidin-GBP-PA. Vector and other CDSs are the  
 same as pPA-GBP and pStreptavidin-GBP.

<220>  
 <221> CDS  
 <222> (1)..(1173)  
 <223> CDS for fusion protein His6-streptavidin-GBP-PA consisting of 391  
 amino acid residues.

<400> 6	
atg aga gga tcg cat cac cat cac cat cac gga tcc ggt tct ggt ggc	48
Met Arg Gly Ser His His His His His His Gly Ser Gly Ser Gly Gly	
1 5 10 15	
cat atg gct gaa gct ggt atc acc ggc acc tgg tac aac cag ctg gga	96
His Met Ala Glu Ala Gly Ile Thr Gly Thr Trp Tyr Asn Gln Leu Gly	
20 25 30	
tcc acc ttc atc gtt acc gct ggt gct gac ggt gct ctg acc ggt acc	144
Ser Thr Phe Ile Val Thr Ala Gly Ala Asp Gly Ala Leu Thr Gly Thr	
35 40 45	
tac gaa tcc gct gtt ggt aac gct gaa tct aga tac gtt ctg acc ggt	192
Tyr Glu Ser Ala Val Gly Asn Ala Glu Ser Arg Tyr Val Leu Thr Gly	



50	55	60	
cgt tac gac tcc gct ccg gct acc gac ggt tcc gga acc gct ctg ggt Arg Tyr Asp Ser Ala Pro Ala Thr Asp Gly Ser Gly Thr Ala Leu Gly 65 70 75 80			240
tgg acc gtt gct tgg aaa aac aac tac cgt aac gct cac tcc gct acc Trp Thr Val Ala Trp Lys Asn Asn Tyr Arg Asn Ala His Ser Ala Thr 85 90 95			288
acc tgg tct ggc cag tac gtt ggt ggt gct gaa gct cgt atc aac acc Thr Trp Ser Gly Gln Tyr Val Gly Gly Ala Glu Ala Arg Ile Asn Thr 100 105 110			336
cag tgg ttg ttg acc tcc ggc acc acc gaa gct aac gcg tgg aaa tcc Gln Trp Leu Leu Thr Ser Gly Thr Thr Glu Ala Asn Ala Trp Lys Ser 115 120 125			384
acc ctg gtt ggt cac gac acc ttc acc aaa gtt tcg agc tct ggt agt Thr Leu Val Gly His Asp Thr Phe Thr Lys Val Ser Ser Ser Gly Ser 130 135 140			432
ggc aat ggt cat atg cat gga aaa act cag gca acc agc ggg act atc Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile 145 150 155 160			480
cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser 165 170 175			528
atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His 180 185 190			576
gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys 195 200 205			624
act cag gca acc agc ggg act atc cag agc atg cat gga aaa act cag Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln 210 215 220			672
gca acc agc ggg act atc cag agc atg cat gga aaa att cag gca acc Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile Gln Ala Thr 225 230 235 240			720
agc ggg act atc cag agc atg cat gct ctg tcc ctc gag ggt ggc gga Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu Gly Gly Gly 245 250 255			768
tcc ggt tct ggt gcg caa cac gat gaa gcc gta gac aac aaa ttc aac Ser Gly Ser Gly Ala Gln His Asp Glu Ala Val Asp Asn Lys Phe Asn 260 265 270			816
aaa gaa caa caa aac gcg ttc tat gag atc tta cat tta cct aac tta Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile Leu His Leu Pro Asn Leu 275 280 285			864

aac gaa gaa caa cga aac gcc ttc atc caa agt tta aaa gat gac cca Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro 290 295 300	912
agc caa agc gct aac ctt tta gca gaa gct aaa aag cta aat gat gct Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala 305 310 315 320	960
cag gcg ccg aaa gta gac aac aaa ttc aac aaa gaa caa caa aac gcg Gln Ala Pro Lys Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala 325 330 335	1008
ttc tat gag atc tta cat tta cct aac tta aac gaa gaa caa cga aac Phe Tyr Glu Ile Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn 340 345 350	1056
gcc ttc atc caa agt tta aaa gat gac cca agc caa agc gct aac ctt Ala Phe Ile Gln Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu 355 360 365	1104
tta gca gaa gct aaa aag cta aat gat gct cag gcg ccg aaa gta gac Leu Ala Glu Ala Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys Val Asp 370 375 380	1152
gcg aat tcg agc tct ggt ggc taa Ala Asn Ser Ser Ser Gly Gly 385 390	1176

<210> 7  
 <211> 393  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <223> CDS portion (115-504) of the expression plasmid, pGBP (total length 5073 base pairs) for fusion protein His6-GBP. Vector and other CDSs are the same as pPA-GBP.

<220>  
 <221> CDS  
 <222> (1)..(390)  
 <223> CDS for fusion protein His6-GBP consisting of 130 amino acid residues.

<400> 7 atg aga gga tcg cat cac cat cac cat cac gga tcc gga ggt ggg agc Met Arg Gly Ser His His His His His His Gly Ser Gly Gly Gly Ser 1 5 10 15	48
tct ggt agt ggc aat ggt cat atg cat gga aaa act cag gca acc agc Ser Gly Ser Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser	96



tct ggt agt ggc aat ggt cat atg cat gga aaa act cag gca acc agc Ser Gly Ser Gly Asn Gly His Met His Gly Lys Thr Gln Ala Thr Ser 20 25 30	96
ggg act atc cag agc atg cat gga aaa act cag gca acc agc ggg act Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr 35 40 45	144
atc cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln 50 55 60	192
agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met 65 70 75 80	240
cat gga aaa act cag gca acc agc ggg act atc cag agc atg cat gga His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly 85 90 95	288
aaa act cag gca acc agc ggg act atc cag agc atg cat gga aaa att Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Ile 100 105 110	336
cag gca acc agc ggg act atc cag agc atg cat gct ctg tcc ctc gag Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Ala Leu Ser Leu Glu 115 120 125	384
ggt ggt gga agc tct ggt agt ggc aat ggt cat atg cat gga aaa act Gly Gly Gly Ser Ser Gly Ser Gly Asn Gly His Met His Gly Lys Thr 130 135 140	432
cag gca acc agc ggg act atc cag agc atg cat gga aaa act cag gca Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala 145 150 155 160	480
acc agc ggg act atc cag agc atg cat gga aaa act cag gca acc agc Thr Ser Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser 165 170 175	528
ggg act atc cag agc atg cat gga aaa act cag gca acc agc ggg act Gly Thr Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr 180 185 190	576
atc cag agc atg cat gga aaa act cag gca acc agc ggg act atc cag Ile Gln Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln 195 200 205	624
agc atg cat gga aaa act cag gca acc agc ggg act atc cag agc atg Ser Met His Gly Lys Thr Gln Ala Thr Ser Gly Thr Ile Gln Ser Met 210 215 220	672
cat gga aaa att cag gca acc agc ggg act atc cag agc atg cat gct His Gly Lys Ile Gln Ala Thr Ser Gly Thr Ile Gln Ser Met His Ala 225 230 235 240	720
ctg tcc ctc gag ggt ccg taa	741

Leu Ser Leu Glu Gly Pro  
245